

## МНОГОЕЗИЧЕН УЕБ-БАЗИРАН РЕДАКТОР ЗА ОПИСАНИЕ НА СЕМАНТИЧНО СТРУКТУРИРАНИ ЗНАНИЯ И ИНФОРМАЦИЯ<sup>1</sup>

**докторант Иван Ангелов**

Пловдивски университет „Паисий Хилендарски“  
Факултет по математика и информатика  
бул. „България“ 236, 4003 Пловдив, България  
iangelov@uni-plovdiv.bg

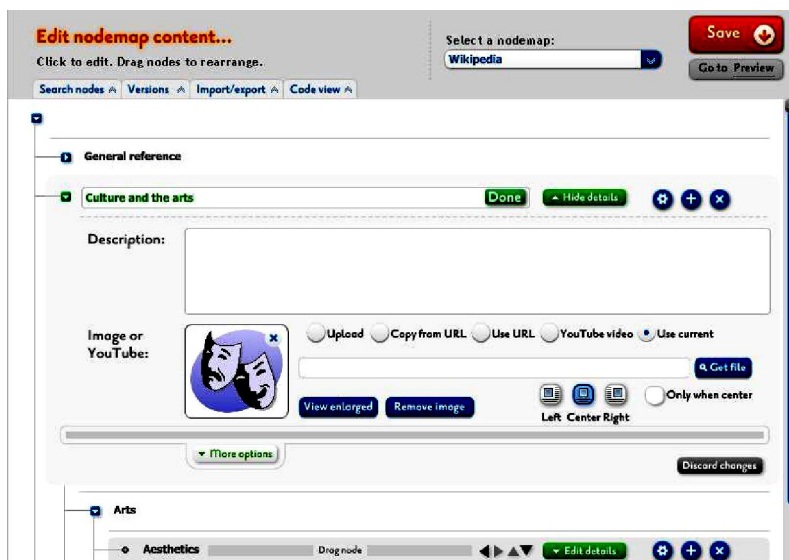
**доц. д-р Асен Рахнев**

Пловдивски Университет „Паисий Хилендарски“  
Факултет по Математика и Информатика  
бул. „България“ 236, 4003 Пловдив, България  
assen@uni-plovdiv.bg

Представен е един интуитивен уеб-базиран редактор на съдържание, който е подходящ за интерактивно описание на йерархично структурирани и семантично свързани знания и информация. Редакторът предоставя многоезичен потребителски интерфейс за въвеждане, форматиране и структуриране на многоезични данни. Заради своя XML-базиран изходен формат той е подходящ за удобно преобразуване на информация. Преобразуването се извършва чрез прости XSL трансформации, към входния работен формат на различни съществуващи продукти, базирани на широк клас представяния. Например: интерактивни радиални карти, хиперболични дървета, карти на сайта и мрежи. Такива биват най-често продукти за графична визуализация на взаимовръзките между отделни категории, концепции или единици знания; уеб-услуги за визуализация или управление на съдържание; продукти за търсене и навигация в структурирана информация и др.

**Ключови думи:** многоезичност, потребителски интерфейс, интернационализация, локализация, превод, персонализиране

**Въведение.** Всеизвестен факт е, че за разбирането и удобната работа на потребителите на софтуерни продукти от съществена важност е потребителския интерфейс да използва познат, а по възможност и майчин език за описание на текстовите полета, менюта, системни съобщения, изскачащите при преминаване под показалеца на мишката подсказващи текстове, помощна информация, синхронизация [7] и т.н.



Фигура 1. Изглед от уеб-редактора на съдържание на SpicyNodes **Error! Reference source not found.** илюстриращ един съвременен, в случая Flash-базиран, потребителски интерфейс от високо равнище.



Фигура 2. Локализиран на български фрагмент от потребителския интерфейс на редактора.

<sup>1</sup>Тази работа е подпомогната по проекти ИС-М-4/2008, РС09-ФМИ-014/17.04.2009 и МУ-6/2008 на поделение „Научна и приложна дейност“ при Пловдивския университет „Паисий Хилендарски“.

От друга страна, разбира се стои и въпросът с многоезичното съдържание, което най-често този интерфейс има за цел да представя и обработва, затова ще се опитаме да насочим тук предлагания подход към многоезичността, както на интерфейса, така и на съдържанието.

**Многоезичен потребителски интерфейс.** В практиката обаче, често се оказва не съвсем достатъчно да локализираме един интерфейс с даден език или негова географска разновидност, както понастоящем разпространените подходи най-често се ограничават да правят. Широк кръг потребители могат да използват различни терминологии, понякога дори и жаргон, за означаване на семантично еднакви понятия или просто бидейки с различен образователен профил или обучени по различни учебни и фирмени програми, да подразбират различни значения или контекст.

С оглед подобряването на потребителския интерфейс в това отношение, целящо максимално интуитивна работа на потребителите с интерфейса, се оказва целесъобразно да преминем отвъд границата на локализирането и интернационализирането на софтуерните продукти, и да погледнем също към профилирането и персонализирането на потребителския им интерфейс. В практиката това би означавало изграждане и предоставяне на възможности в продукта, или на неговия потребителски интерфейс в частност, потребителите самостоятелно да променят, превеждат или дописват текстовите етикети, менюта и помощна информация.

На базово ниво това може да бъде специализиран подход предназначен за обучени потребители, супер-потребители отговарящи например за група от потребители или администратори на конкретната интеграция на продукта.

Ще покажем два основни подхода за реализиране на такава функционалност, единият по-консервативен и използващ гъвкав и стандартен формат; и другият по-нестандартен, но и по-мощен като възможности и реализация, макар и базиран на прост текстов формат, което от своя страна предполага още едно предимство: позволява по-лесна редакция от самия потребител със стандартни и достъпни средства.

Най-честата реализация на така описаната функционална поддръжка на многоезичност включва употребата на хеш-ключове, обикновено именувани на основния език, или още познат като език по подразбиране, на интерфейса. Тези хеш-ключове впоследствие се използват за намиране на локализиран аналог на желаното съдържание, тъй като аналогични съдържания в различни локализации имат еднакъв ключ. Конкретното съдържание от тези кандидати с търсения ключ се определя в зависимост от текущите регионални настройки на операционната система, програмния продукт, потребителските предпочитания или комбинация от трите.

- XML използващ CDATA за локализирани текстове

Пример:

```
<label id="Culture and the arts">
    <![CDATA[ Culture and the arts ]]>
</label>
```

Употребата на CDATA секцията в XML носи допълнителното предимство от улеснена интеграция и реализация на евентуална защита на съдържанието [2] чрез шифриране, подписване или друга DRM [3] имплементация, поради предоставяната възможност за вмъкване на двоични, а не само на текстови данни в съдържанието.

Основното предимство на този подход е неговата висока стандартизираност и широката му употреба и поддръжка в множество понастоящем разпространени продукти за локализация или среди за разработка на потребителски интерфейси. За съжаление обаче, XML не ни дава възможност да задаваме тагове като:

```
<Culture and the arts>
```

което би ни позволило да използваме XML формата за реализиране на подобна гъвкавост, както при следващия предлаган подход, а именно възможността ключ към локализация на съдържанието да бъде самото съдържание по подразбиране.

- Текстови файлове базирани на Unicode [4] таблици на символите

Пример:

файл *Default.properties*:

```
Test=Test
month.1=Jan
Demo=Demo
Trial="a test iteration"
.....
```

файл *bg\_BG.properties*:

```
Test=Тест
```

Demo=Демо  
month.1=Януари

.....  
файл *username.properties*:  
Test=Пробване на промените  
month.1=Ян.  
.....

В горния пример използваме три текстови файла с име Default (по подразбиране), bg\_BG (класификационен код на локализацията) и потребителското име.

Употребата на Unicode кодиране на текстовия файл е обикновено стандартно и включено по подразбиране в съвременните операционни системи. То осигурява възможността в един файл да бъдат употребявани различни езици и ориентации без последваща загуба на локализираща информация.

При нужда на потребителския интерфейс за изобразяване на конкретно локализирано съдържание или подсказваща информация, той се обръща към често наричания „преводач на съдържание“. Преводачът от своя страна започва обхождане на файловете, които по същество представляват именувано множество от двойки „ключ-съдържание“, в ред от най-конкретните локализации към най-общите. В нашия случай това са *username.properties* -> *bg\_BG.properties* -> *Default.properties*. Двойките стойности в тези файлове съдържат примери за типични ключове (*month.1*) и за съдържание като ключ (*Test, Демо, Trial, ..*).

Така при заявка за: „*month.1*“ ще се избере „Ян.“ като най-специфичен, в случая зададен като предпочитание от активния потребител, примерно за краткост или по-малко място на екрана.

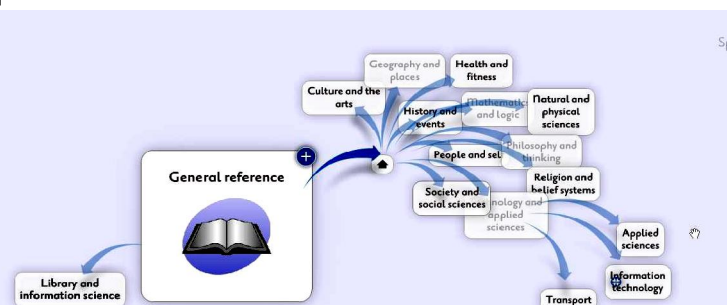
За „*Demo*“ -> „Демо“

За „*Trial*“ -> „a test iteration“

За „*Experiment*“ -> „*Experiment*“ (самият ключ се подразбира за съдържание, когато не се намери съвпадение, в противен случай се избира най-специфичното, т.е. първото открито съответствие).

Важно за отбелязване е също и илюстрираната възможност потребителя сам да конкретизира, разяснява или допълва понятия, които имат неясен или двусмислен за него контекст (Вж. Test).

**Многоезично съдържание.** В повечето продукти използващи потребителски интерфейс, очаквано намираме и възможност за задаване, обработка или друга употреба на дефинирано от потребителя съдържание като текстови полета за набор или изграждане на структурирани документи за по-нататъшна обработка и с различно предназначение. Това съдържание и тези документи, задавани или въвеждани от широк кръг потребители, от своя страна също подлежат на локализиране, превод и/или интернационализация.



Фигура 3. Визуализация от веб-услугата SpicyNodes на един примерен изходен документ, описан и локализиран от редактора.

За да демонстрираме по-пълно подхода в неговата цялост разглеждаме една примерна реализация на редактор на съдържание, при това с технически по-предизвикателния, веб-базиран потребителски интерфейс. Фактът, че редактора използва потребителски интерфейс от високо равнище, неговото предназначение за работа с глобален кръг потребители, и поради това силно интернационализирано, съдържание преценихме, че това е един изчерпателен пример на разглеждания подход за референтна имплементация. Пълнотата и изчерпателността на примерната практическа реализация се обуславя допълнително и от избора на самия редактор, на който се спряхме, а именно веб-редактора на SpicyNodes, позволяващ въвеждане, редактиране, подредба и структуриране на семантично свързано съдържание, при това с изход в един отворен и гъвкав формат приложим за обработка и визуализация от различни други софтуерни продукти, от различни външни разработки и програми осигуряване.

Възможностите за форматиран изход на този редактор, с цел обработката на редактираното съдържание с други програмни продукти или архивирането му, са три:

1. Собствен вътрешен текстово-базиран формат;
2. XML-базиран собствен формат за обмен на съдържание предимно с модула за визуализация и разглеждане на SpicyNodes

Пример:

```
<node propagate="true" imgalign="center">
  <label><![CDATA[ Culture and the arts ]]></label>
  <node propagate="true" imgalign="center">
    <label><![CDATA[ Language ]]></label>
  </node>
</node>
```

3. XML-базиран формат съвместим с получилия разпространение формат на FreeMind, софтуер за изграждане и визуализация на ментални карти [5].

Първият от горните, вътрешен текстов формат е удобен и предназначен за потребители без познания за XML или не разполагачи с удобен за целта редактор, но същевременно желаещи или имащи нужда от фина настройка на атрибутите на визуализация и представяне на информацията, включително нейната структура, да сторят това без специални знания или софтуерни инструменти, редактирайки един обикновен текстов файл.

Другите два, по същество XML-базирани формати, имат предимството да бъдат лесно разчитани и разпознавани от софтуерни продукти на други разработчици или от човешки потребители и редактори с познания по XML. Те са също така лесни за преобразуване към други несъвместими, но все пак структурирани предимно йерархично, формати чрез например употребата на прости XSL трансформации [6].

**Заклучение.** Нашето мнение е, че е целесъобразно да се реализира тази функционалност и да се насочи към още по-широк кръг от потенциални потребители, т.е. на всеки един потребител да бъде предоставена възможността да си превежда, или дописва, един по-общ превод, или пък да преименува непознато звучащи му термини с познати и интуитивни такива. Още по-добре би било, тази функционалност да бъде предоставена по време на самата работа на потребителя с продукта или неговия интерфейс, например закачането на контекстно меню или събитие генерирано от мишката, клавишна комбинация или друго устройство за взаимодействие с човешки потребител, на специализиран диалог предоставящ възможност за превод, промяна или преименуване на текущия, или друг указан, елемент от потребителския интерфейс и неговите текстови атрибути.

Всичко това може да се случва динамично, в реално време и с резултат в сила веднага след въвеждането на промените в специализирания диалог. Понякога, а може би по-скоро често, с оглед на по-висока безопасност и защита от случайна промяна от потребителите, това се препоръчва да се прави само в специален тестов режим на работа на продукта. В същото време промените ще бъдат и постоянно запазени в настройките или постоянната памет на устройството или конкретния продукт за поддържащата им се употреба в следващите сесии на потребителя с интерфейса.

Желателно, и силно препоръчително, е така описаните мерки и подходи да бъдат реализирани, прилагани и употребявани в среди за описание и управление на потребителски интерфейси, които предлагат и поддържат мениджъри за гъвкаво и динамично разполагане на елементите на потребителския интерфейс, или да бъде обърнато специално внимание на тези изисквания, с оглед запазването на неговата цялост, общ изглед и последователен вид, за да не бъдат те нарушавани най-вече от динамична промяна в дължината и/или височината на елемента, или дори неговата ориентация по време на работа на продукта, например при поискване по гореописания начин на една евентуална незабавна локализация.

#### ЛИТЕРАТУРА:

- [1] Douma, M., Ligierko, G., Angelov, I. SpicyNodes Radial Map Engine – In Proceedings of the 34th Conference on Applications of Mathematics in Engineering and Economics (AMEE '08) - June 8-14, Sozopol, Bulgaria, ISBN: 978-0-7354-0598-1, vol. 1067, pp. 209-220
- [2] Ангелов, И. – Защита на Учебно Съдържание чрез Уеб-управление на Потребителските Права – Сборник на Международната конференция "Компютърни методи в науката и образованието", 12 – 14 Септември, 2008, Варна – под печат
- [3] Kumik, P. Digital Rights Management. Legal Information Management, 1, 2001, pp 21-23, doi:10.1017/S147266960000402
- [4] Aliprand, J., Allen, J., Becker, J., Davis, M., Everson, M., Freytag, A., Jenkins, J. H., Ksar, M., McGowan, R., Moore, L., Suignard, M., and Whistler, K. 2000 The Unicode Standard Version 3.0. Addison Wesley Longman Publishing Co., Inc., ISBN:0-201-61633-5
- [5] Polivaev, D. FreeMind – freemind mapping software, FreeMind Official Homepage & Wiki, Available at [http://freemind.sourceforge.net/wiki/index.php/Main\\_Page](http://freemind.sourceforge.net/wiki/index.php/Main_Page), 2008.
- [6] Clark, J. XSL Transformations (XSLT) Version 1.0, W3C Recommendation, <http://www.w3.org/TR/xslt> (November, 1999).
- [7] Valchanov, N., T. Terzieva, A. Iliev. A model of a system for centralized synchronization of independent applications, 6th Mediterranean Conference on Mathematics Education, Plovdiv 22-26.4.2009, 459-466.